

# CS448 PSO

## Week 4

CS448 staff

### 3NF & BCNF

- FD:  $X \rightarrow A$  is “Nontrivial” = right-side attribute not in left side
- 3NF: A relation  $R$  is in 3NF iff (if and only if)

for every nontrivial FD:  $X \rightarrow A$ , either:

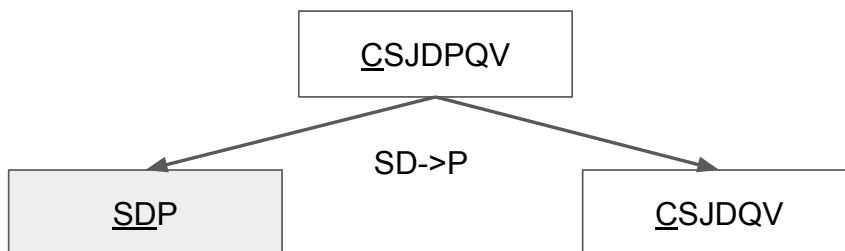
- $X$  is a superkey, or
- $A$  is prime = member of at least one key.
- A relation  $R$  is in BCNF if for every nontrivial FD for  $R$ , say  $X \rightarrow A$ ,
  - $X$  is a superkey
  - ~~$A$  is prime = member of at least one key.~~

## Example 1: decompose a schema into BCNF

- Schema: CSJDPQV
- FDs: {SD- $\rightarrow$ P, J- $\rightarrow$ S, JP- $\rightarrow$ C}
- FDs violate BCNF

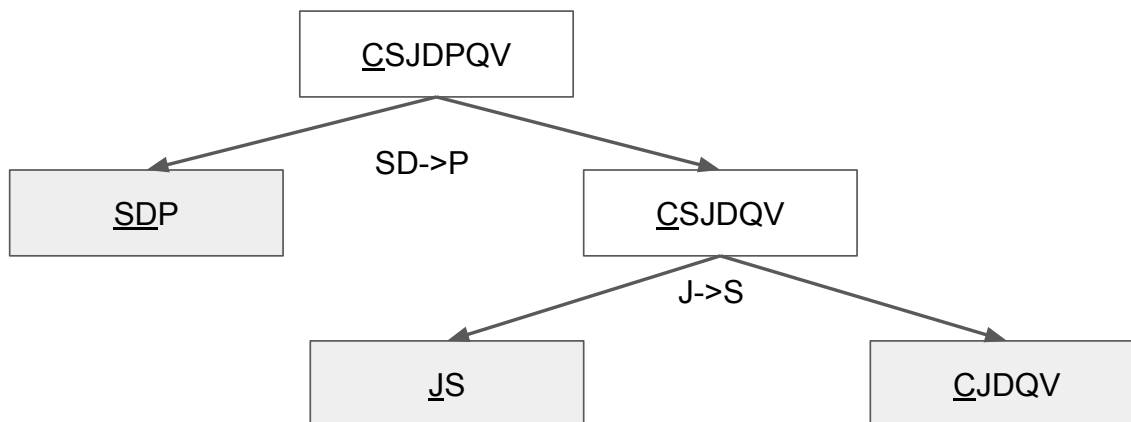
## Example 1: decompose a schema into BCNF

- Schema: CSJDPQV
- FDs: {SD- $\rightarrow$ P, J- $\rightarrow$ S, JP- $\rightarrow$ C}
- FDs violate BCNF



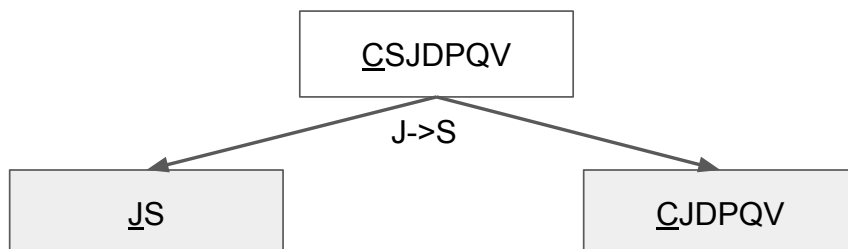
## Example 1: decompose a schema into BCNF

- Schema:  $\underline{C}SJDPQV$
- FDs:  $\{SD \rightarrow P, J \rightarrow S, JP \rightarrow C\}$
- FDs violate BCNF



## Alternatives in decomposing into BCNF

- Schema:  $\underline{C}SJDPQV$
- FDs:  $\{SD \rightarrow P, J \rightarrow S, JP \rightarrow C\}$
- FDs violate BCNF

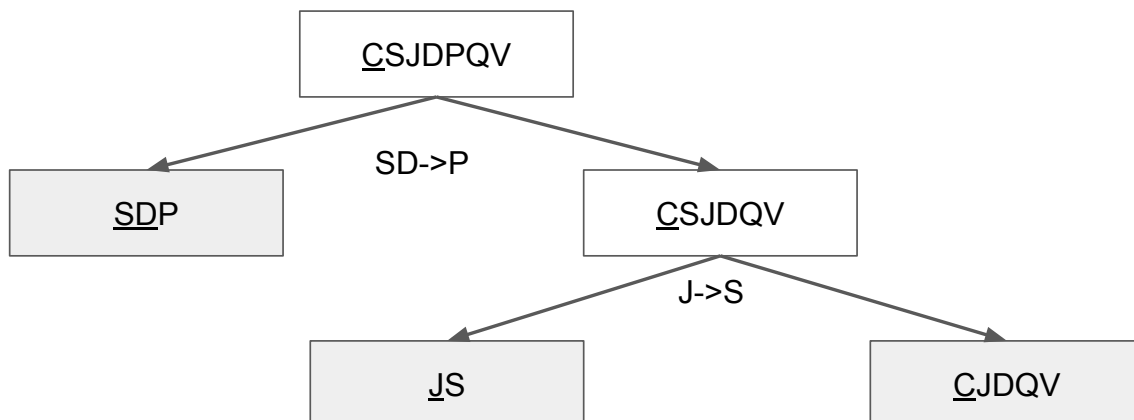


## Which alternative should be used?

- Choose the alternatives based on the semantics of the application.
- Example:  $R = (\text{course id}, \text{course name}, \text{course abbreviation}, \text{year}, \text{instructor})$ 
  - $\text{course abbreviation} \rightarrow \text{course name}$
  - $\text{course name}, \text{year} \rightarrow \text{instructor}$
- The most frequently used query is selecting instructors given the course name and year.
- Two decompositions:
  - $(\text{course name}, \text{course abbreviation})$  and  $(\text{course id}, \text{course abbreviation}, \text{year}, \text{instructor})$
  - $(\text{course name}, \text{year}, \text{instructor})$  and  $(\text{course id}, \text{course name}, \text{course abbreviation}, \text{year})$

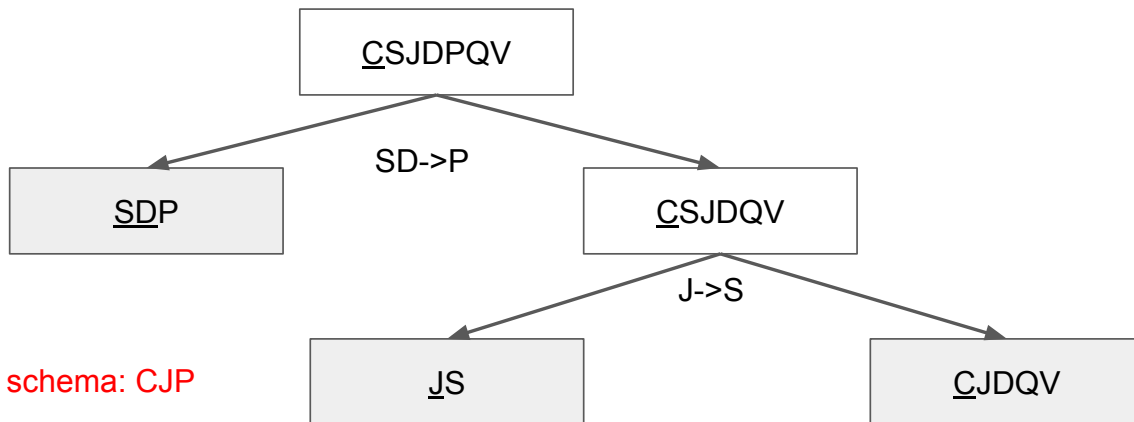
## Example 2: dependency-preserving decomposition into 3NF

- Schema:  $\underline{C}SJDPQV$
- FDs:  $\{SD \rightarrow P, J \rightarrow S, \textcolor{red}{JP} \rightarrow \textcolor{red}{C}\}$
- $SD \rightarrow P, J \rightarrow S$  violate 3NF



## Example 2: dependency-preserving decomposition into 3NF

- Schema:  $\underline{C}SJDPQV$
- FDs:  $\{SD \rightarrow P, J \rightarrow S, JP \rightarrow C\}$
- $SD \rightarrow P, J \rightarrow S$  violate 3NF



## Create views to consolidate a non-preserved FD

- Use materialized views to consolidate a non-preserved FD into one table
- Check the FD in that materialized view by making LHS of the FD the key for the view
- But will need to maintain the views when the base tables get updated

## 3NF synthesis

Take all the attributes over the original relation  $R$  and a minimal cover  $F$  for the FDs that hold over it and add a relation schema  $XA$  to the decomposition of  $R$  for each FD  $X \rightarrow A$  in  $F$ .

If no decomposed table contains a candidate key for  $R$ , add a relation schema of any candidate key for  $R$ .

## Minimal cover

Minimal (Canonical) Cover for a set  $F$  of FDs is a set  $G$  of FDs such that:

1. Every dependency in  $G$  is of the form  $X \rightarrow A$ , where  $A$  is a single attribute
2. The closure  $F^+$  is equal to the closure  $G^+$
3. If we obtain a set  $H$  of dependencies from  $G$  by deleting one or more dependencies or by deleting attributes from a dependency in  $G$ , then  $F^+ \neq H^+$

## Example 3: find the minimal cover set

- $A \rightarrow B, ABCD \rightarrow E, EF \rightarrow G, EF \rightarrow H, ACDF \rightarrow EG$
- 1. Rewrite  $ACDF \rightarrow EG$  to  $ACDF \rightarrow E, ACDF \rightarrow G$
- 2.  $ACDF \rightarrow G$  is redundant
- 3.  $ACDF \rightarrow E$  is redundant
- 4. Minimize left side of  $ABCD \rightarrow E$  to  $ACD \rightarrow E$
- $A \rightarrow B, ACD \rightarrow E, EF \rightarrow G, EF \rightarrow H$

## General algorithm for obtaining a minimal cover set

1. Put the FDs in a standard form (single attribute on the right)
2. Minimize the left side of each FD
3. Delete redundant FDs

## Example 4: 3NF synthesis

- Schema: CSJDPQV
  - FDs: {SD-→P, J-→S, JP-→C, C-→CSJDPQV}
1. Find the minimal cover set

## Example 4: 3NF synthesis

- Schema: CSJDPQV
  - FDs: {SD-→P, J-→S, JP-→C, C-→CSJDPQV}
1. Find the minimal cover set

C-→S, C-→J, C-→D, C-→P, C-→Q, C-→V, SD-→P, J-→S, JP-→C



## Example 4: 3NF synthesis

- Schema: CSJDPQV
- FDs: {SD-→P, J-→S, JP-→C, C-→CSJDPQV}

1. Find the minimal cover set

~~C~~→~~S~~, ~~C~~->~~J~~, ~~C~~->~~D~~, ~~C~~→~~P~~, C-→Q, C-→V, SD-→P, J-→S, JP-→C

## Example 4: 3NF synthesis

- Schema: CSJDPQV
- FDs: {SD-→P, J-→S, JP-→C, C-→CSJDPQV}

1. Find the minimal cover set

~~C~~→~~S~~, C-→J, C-→D, ~~C~~→~~P~~, C-→Q, C-→V, SD-→P, J-→S, JP-→C

2. Schemas: CJ, CD, CQ, CV, SDP, JS, JPC
3. (Optional) Combine: CJDQVP, SDP, JS
4. CJDQVP is a superkey
5. Done

## Case Study: The Internet shop

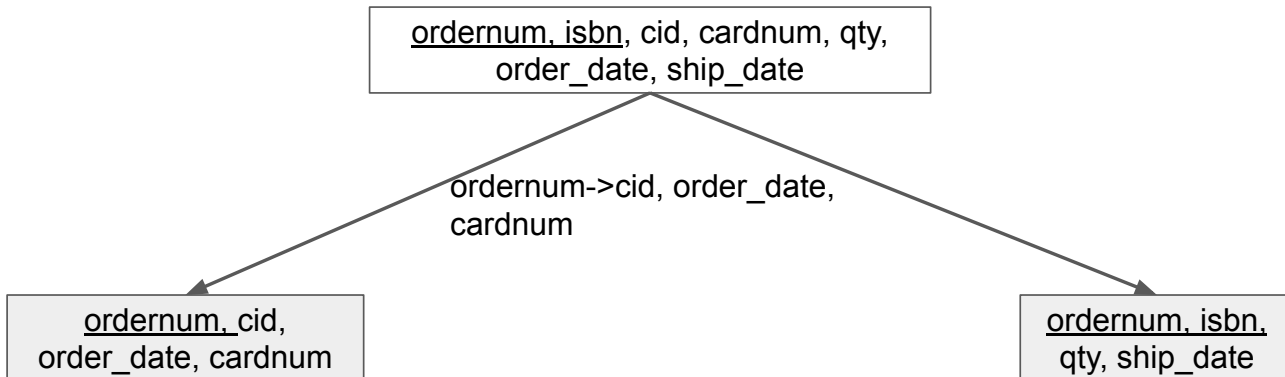
- Books (isbn: VARCHAR(10), title: VARCHAR(8), author: VARCHAR(80), qty\_in\_stock: int, price: double, year\_published: int)
- Customers (cid: int, cname: VARCHAR(80), address: VARCHAR(200))
- Orders (ordernum: int, isbn: VARCHAR(10), cid: int, cardnum: VARCHAR(16), qty: int, order\_date: DATE, ship\_date: DATE)

## Functional dependencies

- Books has one key: isbn. No other FDs
- Customers has one key: cid. No other FDs
- Orders has key: (ordernum, isbn). ordernum->cid, ordernum->order\_date, ordernum->cardnum

# Decomposition

- Orders (ordernum: int, isbn: VARCHAR(10), cid: int, cardnum: VARCHAR(16), qty: int, order\_date: DATE, ship\_date: DATE)
- ordernum->cid, ordernum->order\_date, ordernum->cardnum



# Decomposition

- Orders (ordernum: int, isbn: VARCHAR(10), cid: int, cardnum: VARCHAR(16), qty: int, order\_date: DATA, ship\_date: DATE)
- ordernum->cid, ordernum->order\_date, ordernum->cardnum
- Orders (ordernum, cid, order\_date, cardnum)
- Orderlists (ordernum, isbn, qty, ship\_date)

# References

1. "The cow book": Database management systems by Raghu Ramakrishnan and Johannes Gehrke

